Abundance estimates for Antarctic minke whales using the OK method

Hiroshi Okamura[1] and Toshihide Kitakado[2]

[1] National Research Institute of Far Seas Fisheries, Fisheries Research Agency, Kanagawa 236-8648, Japan

[2] Tokyo University of Marine Science and Technology, Minato, Tokyo 108-8477, Japan

The corresponding author's email address: okamura@fra.affrc.go.jp

ABSTRACT

The abundance estimates of Antarctic minke whales using the OK method are presented with some diagnostics. The OK method is a design-based hazard probability model. It was applied to the preferred dataset and the reference dataset for comparison. The total abundances for the preferred dataset were 1,405,264 (CV: 0.118) for CPII and 674,128 (CV: 0.088) for CPIII with the actual northern boundary of the surveyed strata, when we used the definite duplicates only. Those for the reference dataset were 1,350,210 (CV: 0.122) for CPII and 647,227 (CV: 0.100) for CPIII. When we used the similar specification to SPLINTR, the abundances were 1,179,581 (CV: 0.116) for CPII and 631,915 (CV: 0.098) for CPIII. As a whole, the OK method provided robust estimates for different datasets and specifications. CPII estimates were more sensitive to different datasets and assumptions than CPIII estimates.

1. INTRODUCTION

There are two available approaches for estimating abundances from IDCR/SOWER data taking $g(0)$ estimation into account: One is SPLINTR (Bravington and Hedley 2010) which is a spatial model and assumes trackline independence for duplicate sightings, another is OK (Okamura and Kitakado 2010) which is a design-based model and uses a hazard probability approach (e.g., Schweder et al. 1997). The abundance estimates from two methods were considerably different (IWC 2010). The OK method provided about double circumpolar estimates compared to those from the standard method (Branch 2006), while the estimates from the SPLNTR method were almost the same level of Branch's estimates. The diagnostics were generally good for both models. Bravington and Hedley (2010) developed non-spatial (stratified) SPLINTR model which removed spatial components from the original SPLNTR model (non-stratified SPLNTR). The difference between stratified SPLNTR and non-stratified SPLNTR was less than 10% in terms of abundance estimates and it was insufficient to explain the substantial difference between OK and SPLNTR. In the 2010 IWC/SC meeting, the fact that the difference between two models is attributable to the difference of effective strip width (esw) was confirmed. In the Bergen workshop held this January (IWC 2011), Okamura and Kitakado (2011) and Bravington (2011) examined the effect of using the trackline independence (TLI) assumption. Their results indicated that the TLI assumption would bring about ca. 20% to 25% reduction of abundance compared to hazard probability models. During the Bergen workshop, some other potential factors that might produce difference were identified and agreed.

Because it was considered that only one factor could not explain the difference between two methods in the Bergen workshop, using the step-by-step sensitivity analyses to be compared between them was agreed. We identified necessary sensitivity analyses in the Bergen workshop (IWC 2011).

The data used by each modeler are different due to the specifications of models and pre-processing of data. Therefore, the "reference" dataset was produced through intersessional email correspondences. Although there was a minor difference of "processed" reference datasets between SPLINTR and OK, it was agreed between the modelers that the difference is small and does not produce any serious problems.

The workshop agreed that the modelers could use two datasets, "reference" and "preference". The reference dataset is just for comparison. The preferred dataset is the dataset that each modeler thinks reasonable to their model. Sensitivity runs are conducted based on the reference dataset. The OK method uses the confirmation information about observed school sizes while the SPLINTR uses all Independent Observer (IO) mode sightings as unconfirmed and all Closing (CL) mode sightings as confirmed. Furthermore, SPLINTR uses School size experiment (SSX) data and OK does not. Thus, the OK method was proposed to change treatment of confirmation and SSX data in a step-by-step way to imitate the specification of SPLNTR. In addition, some additional sensitivity analyses such as including possible duplicates and special handling of sightings with school size = 1 were also proposed.

This document presents the specifications of the latest OK method and the results of fitting the model to some variants of the IDCR/SOWER datasets as sensitivity analyses according to the proposal in the Bergen workshop (IWC 2011).

2. MATERIALS AND METHODS

2.1. The data

The details of our preferred dataset are given in Appendix A. The basic differences between the reference dataset and the preferred dataset are: 1) use of 1992 data for estimating detection function parameters, 2) the effort of the preferred dataset is larger by 13%, and 3) the sighting number of the preferred dataset is larger by 13%. About 1), there was no special reason to remove 1992 data in estimating detection function parameters. (2) and (3) means that some data would have been removed from the preferred data in the reference data. Because clear grounds for removal are unknown to us, we prefer keeping the data.

2.2. The OK method

The details of the OK method are given in Appendices B to F. It is based on a hazard probability model (Schweder et al. 1997) and the variance of abundance estimates are calculated by a design-based estimator. The measurement errors are taken into account in the likelihood function (Appendix C).

2.3. Sensitivity analyses

Abundance estimates are calculated according to the procedure given in Appendices B to F for both the preferred and reference datasets. For the preferred dataset, we also calculate abundance estimates: 1) when the "BH" code is included, and 2) when the possible duplicates are included. When the "BH" code is included, only the sightings by topmen are used when estimating detection function parameters and density for the sightings under the "BH". For the reference dataset, first all IO mode data are transformed to unconfirmed and the model is fitted. Then, in addition to previous change, all CL mode data are transformed to confirmed and the model is fitted. Finally, the probability function for the SSX data are added to the likelihood in addition to IO = unconfirmed and CL = confirmed. As other sensitivity analyses, for the reference dataset, we calculate abundance estimates: 1) when the school size error by unconfirmation is set to zero, and 2) when parameterization for school size 1 sightings are treated separately.

3. RESULTS AND DISCUSSION

To save the space, the diagnostics was not given in this paper. The diagnostics for all the runs were generally good. By request and in the small group meeting, all diagnostics will be provided from the first author.

Table 1 provides the abundance estimates for the preferred and reference datasets using the survey once method (Branch and Butterworth 2006), where the run for the reference dataset was changed to approach the specification of SPLNTR in a step-by-step way. The total abundances for the preferred dataset were 1,405,264 (CV: 0.118) for CPII and 674,128 (CV: 0.088) for CPIII. The absolute values of abundance estimates for the preferred dataset were the highest. On the other hand, the ratio of CPII and CPIII abundance estimates increased in a step-by-step way. The abundance estimates for the reference dataset were 1,350,210 (CV: 0.122) for CPII and 647,227 (CV: 0.100) for CPIII. When we used SSX data in addition to the change of definition of confirmation status, the abundances were 1,179,581 (CV: 0.116) for CPII and 631,915 (CV: 0.098) for CPIII. The changes of CPII were greater than those of CPIII. When all IO data were set to unconfirmed, the influence was the largest. Conversely, setting all CL data to confirmed made abundance estimates higher. Setting only IO to unconfirmed would bring about underestimation of school size bias effect because confirmed sightings that do not have bias in reality are regarded as unconfirmed. Setting CL mode data to confirmed offsets such underestimation. Therefore, the results for changes of confirmation status are intuitively understandable. The use of SSX data lowers abundance estimates slightly. CPII abundance estimates are more sensitive to the change of data and assumption than CPIII abundance estimates. Additional variance was not added to any estimate. It will be calculated during the meeting.

Table 2 provides the abundance estimates for the additional sensitivity analyses. When the school size error for unconfirmed sightings was set to zero, the abundance decreased considerably, especially CPII did by about half. Including the "BH" code did not change the abundance estimates very much (CPII does not include "BH" code). Setting different parameters for school size one increased abundance estimates because esw decreased for school size one sightings. Including possible duplicates decreased abundance estimates as expected. In general, CPII is more sensitive to the changes given in Table 2.

REFERENCES

Branch, T. A. (2006) Abundance estimates for Antarctic minke whales from three completed circumpolar sets of surveys, 1978/79 to 2003/04. Paper SC/58/IA18 presented to the IWC/SC (unpublished): 28pp.

Branch, T. A. and Butterworth, D. S. (2001) Southern Hemisphere minke whales: standardised abundance estimates from the 1978/79 to 1997/98 IDCR-SOWER surveys. J. Cetacean Res. Manage.

3: 143-174.

Branch, T. A. and Butterworth, D. S. (2006) Suggested options for the analysis of IDCR/SOWER minke whale data. Paper SC/58/IA19 presented to the IWC/SC (unpublished): 10pp.

Cooke, J. G. (1997) An implementation of a surfacing-based approach to abundance estimation of minke whales from shipborne surveys. Rep. int. Whal. Commn **47**: 513-528.

Cooke, J. G. (2001) A modification of the radial distance method for dual-platform line transect analysis to improve robustness. Paper SC/53/IA31 presented to the IWC/SC (unpublished): 7pp.

International Whaling Commission (2002) Report of the sub-committee on the comprehensive assessmnet of whale stocks - in-depth assessments. J. Cetacean Res. Manage. **4** (Suppl.): 192-229.

International Whaling Commission (2008) Report of the SOWER Abundance Workshop.

Kishino, H. and Kasamatsu, F. (1987) Comparison of the closing and passing mode procedures used in sighting surveys. Rep. int. Whal. Commn **37**: 253-258.

Kitakado, T. and Okamura, H. (2005) Estimation methods of the additional variance for Antarctic minke whales. Paper SC/57/IA5 presented to the IWC/SC (unpublished): 7pp.

Kitakado, T. and Okamura, H. (2008) Some issues on the additional variance. Paper SC/F08/A11 presented to the SOWER Abundance Workshop (unpublished): 5pp.

Kitakado, T. and Okamura, H. (2009) Estimation of additional variance for Antarctic minke whales based on the abundance estimates from the revised OK method. Paper SC/61/IA8 presented to the IWC/SC meeting (unpublished): 11pp.

Matsuoka, K., Ensor, P., Hakamada, T., Shimada, H., Nishiwaki, S., Kasamatsu, F. and Kato, H. 2003. Overview of minke whale sightings surveys conducted on IWC/IDCR and SOWER Antarctic cruises from 1978/79 to 2000/01. J. Cetacean Res. Manage. 5:173-201.

Monien, H. (2006) Gaussian summation: an exponentially convergent summation scheme. arXiv:math/0611057v1. Available at http://arxiv.org/PS cache/math/pdf/0611/0611057v1.pdf.

Okamura, H., Kitakado, T., Hiramatsu, K., and Mori, M. (2003) Abundance estimation of diving animals by the double-platform line transect method. Biometrics **59**: 512-520.

Okamura, H., Kitakado, T., and Mori, M. (2005) An improved method for line transect sampling in Antarctic minke whale surveys. J. Cetacean Res. Manage. **7**: 97-106.

Okamura, H. and Kitakado, T. (2010) SC/ AE3.

Okamura, H. and Kitakado, T. (2011) Comparison between a point independence model and a hazard probability model using the IDCR/SOWER Antarctic minke whales data. SC/ AE3. (unpblished manuscript)

Schweder, T., Skaug, H. J., Dimakos, X. K., Langaas, M., and Oien, N. (1997). Abundance of northeastern Atlantic minke whales, estimates for 1989 and 1995. Rep. int. Whal. Commn **47**: 453-483.

Appendix A The preferred data

The preferred dataset was constructed in the following ways:

Circumpolar set

We used 1985/86 - 2003/04 data, which corresponded to CPII and CPIII.

Vessel Speed

The vessel speeds recorded in the effort records were used for calculating the traveled distances. When the value was 888 (variable speed) or 999 (missing), we used the preset speed, 12 knot (before 1986/87) and 11.5 knot (after 1987/88). When estimating the detection function parameters, the preset vessel speed was used.

Survey effort

The survey effort was calculated by the vessel speed times the traveled time in the effort records. We used all the data with different activity codes except the code "BH". "BH" was taken into account in a sensitivity analysis.

Species

We used data with the species code 04, 91, 92, 39. "Like minke" was included.

Sea state

We used a new category for simplicity in that 0-3 is good (0) and 4-5 is bad (1).

Platform

The original category is the following: 1 - topman in standard barrel, 2 - topman in IO position, 3 - upper bridge, primary observer, 4 - upper bridge, not primary observer, 5 - 1 and 4 simultaneously, 6 - 2 and 4 simultaneously.

We used a new category in that 1 & 5 → A, 2 & 6 → B, and 3 & 4 → C as in Appendix C.

Sighting distances and angles

Bias-corrected distances and angles were used. Angles were truncated at 90 degrees and transformed to radian. We used the perpendicular and forward distances transformed from the radial distances and angles in the analysis. The perpendicular distances transformed were truncated at 1.5 nautical miles, while the forward distances were not truncated.

School size

Best school size estimates were used.

Duplicate

Duplicate sightings in the IO-tracking searching under closing mode in 1987/88 were removed.

We adopted "definite" duplicates as the true duplicates under the IO mode. However, we also used "definite + possible" duplicates as a sensitivity test. When any covariate other than sighting distances and angles was different in a duplicate sighting, we conformed to the following rules:

- School size: If confirmed school size was only one, we used the value. If there were multiple confirmed school sizes or no confirmed school size, we then used the value of the platform with the highest sighting position based on the notion that a topman was the most reliable.

- Confirmation status of school size: If we had at least one confirmed school size, it was defined as confirmed.
- Sea state: When the sea states were different, we adopted the sea state with the earlier record time.

The case with different covariates by different platforms in a duplicate sighting was few and therefore the above minor adjustment will have little effect on abundance estimation.

Truncation

Perpendicular distances were truncated at 1.5 nautical miles according to conventional method (Branch and Butterworth, 2001). When the sightings were duplicates, we usec the averaged distances for the simultaneous duplicates and the distances of later sightings for the delayed duplicates.

Appendix B Detection probability function of sighting cues

The hazard probability model is given by a logistic form,

$$Q(x, y) = \frac{1}{1 + \exp[\tau_r R^{\gamma_r} + \tau_a A^{\gamma_a} + \omega]} \tag{B.1}$$

where $R = \sqrt{x^2 + y^2}$, $A = \text{atan}(x/y)$, x is the perpendicular distance, y is the forward distance, τ_r, τ_a, γ_r, and γ_a are scalar parameters with positive values. The parameters, τ_r, τ_a, and ω, are related to several covariates through a link function as follows:

$\log(\tau_r) \sim$ Platform + log(School.size) + Weather,

$\log(\tau_a) \sim$ Platform + log(School.size) + Weather,

$\omega \sim$ Platform + log(School.size) + Weather + Vessel.

In addition, the surfacing rate λ in Eq. (1) is modeled to be a function of school size,

$\log(\lambda) \sim$ log(School.size),

where the coefficient of log(School.size) is constrained to be positive.

Appendix C Specification of detection function for each sighting pattern

There are three platforms with two independent observers and one semi-independent observer in the IO mode while there are two platforms with no independent observer in the Closing mode. The detection pattern in the IO mode is therefore complicated by taking account of duplicate sightings.

C-1. IO mode

IO mode has three sighting platforms, the top barrel and the IO booth with independent observers, and the upper bridge with semi-independent observers or researchers. We can have information needed to estimate $g(0)$ from the sighting patterns of independent observers (Schweder et al., 1997; Cooke, 1997; Cooke, 2001; Okamura et al., 2003, 2005). The probability density function for each sighting pattern is given below. The contribution to the likelihood function of detection with each sighting pattern is calculated by each probability density times the probability mass density of school size (Appendix C) divided by $\mathrm{esw}_{A\cup B\cup C}$ when school sizes are confirmed. When school sizes are unconfirmed, the numerator is summed up over all school sizes. $\mathrm{esw}_{A\cup B\cup C}$ is given by

$$\mathrm{esw}_{A\cup B\cup C} = \sum_{s=1}^{\infty} \left[\int_0^{x_{max}} \int_0^{\infty} \frac{\lambda}{v} Q_{A\cup B\cup C}(x, y|s) \right.$$
$$\left. \times \exp \left\{ -\frac{\lambda}{v} \int_y^{\infty} Q_{A\cup B\cup C}(x, y'|s) dy' \right\} dxdy \right] \pi(s), \tag{C.1}$$

which is equal to Eq. (4) when k = IO mode.

We have two distances by independent observers in the delayed duplicates. We use the averaged distances for the simultaneous duplicates and the distances of the latter sightings for the delayed duplicates, since the latter sightings tend to have the distances closer to the vessel which are generally likely to be more accurate. The distances of the first sightings are calculated by adding the vessel speeds times the differences of the recorded times between the two sightings to the distances of the latter sightings.

In the IDCR/SOWER surveys before 1988/89, the sighting time was recorded in a "minute" unit, and "second" was omitted. We therefore add to the model to apply to the data before 1988/89 the additional sturcture taking account of uncertainty by rounding the sighting time to the nearest minute.

1. A

$$p(x, y, A) = \frac{\lambda}{v} \{Q_{A\cup B}(x, y) - Q_B(x, y)\} \exp \left\{ -\frac{\lambda}{v} \int_0^y Q_B(x, y') dy' \right\}$$
$$\times \exp \left[-\frac{\lambda}{v} \left\{ \int_y^{\infty} Q_{A\cup B}(x, y') dy' + \int_{y+vT}^{\infty} Q_{A\cup B\cup C \setminus A\cup B}(x, y') dy' \right\} \right], \tag{C.2}$$

where $T = 90/3600$h (before 1988/89) and $T = 60/3600$h (after 1989/90).

2. B

Same as A except for exchanging the symbols A and B.

3. C

$$p(x, y, C) = \frac{\lambda}{v} \{Q_{A\cup B\cup C}(x, y) - Q_{A\cup B}(x, y)\}$$
$$\times \exp \left[-\frac{\lambda}{v} \left\{ \int_0^y Q_{A\cup B}(x, y') dy' + \int_y^{\infty} Q_{A\cup B\cup C}(x, y') dy' \right\} \right]. \tag{C.3}$$

4. $A \times B$

$$p(x, y, AB) = \frac{\lambda}{v} \left(Q_A(x, y) Q_B(x, y) \exp \left\{ -\frac{\lambda}{v} \int_y^{\infty} Q_{A\cup B}(x, y') dy' \right\} \right.$$
$$+ Q_A(x, y) \exp \left\{ -\frac{\lambda}{v} \int_y^{\infty} Q_A(x, y') dy' \right\}$$
$$\times \left[\exp \left\{ -\frac{\lambda}{v} \int_{y+vT}^{\infty} Q_{A\cup B\setminus A}(x, y') dy' \right\} - \exp \left\{ -\frac{\lambda}{v} \int_y^{\infty} Q_{A\cup B\setminus A}(x, y') dy' \right\} \right]$$
$$+ Q_B(x, y) \exp \left\{ -\frac{\lambda}{v} \int_y^{\infty} Q_B(x, y') dy' \right\}$$
$$\left. \times \left[\exp \left\{ -\frac{\lambda}{v} \int_{y+vT}^{\infty} Q_{A\cup B\setminus B}(x, y') dy' \right\} - \exp \left\{ -\frac{\lambda}{v} \int_y^{\infty} Q_{A\cup B\setminus B}(x, y') dy' \right\} \right] \right) \tag{C.4}$$

where $T = 90/3600$h (before 1988/89) and $T = 60/3600$h (after 1989/90).

5. $A \to B$

For the dataset before 1988/89,

$$p(x,y,A\to B) = \frac{\lambda}{v}$$
$$\times\left[\exp\left\{-\frac{\lambda}{v}\int_{y+v(\tau_{AB}+T)}^{\infty} Q_{A\cup B\setminus B}(x,y')dy'\right\} - \exp\left\{-\frac{\lambda}{v}\int_{y+v(\tau_{AB}-T)}^{\infty} Q_{A\cup B\setminus B}(x,y')dy'\right\}\right]$$
$$\times Q_B(x,y)\exp\left\{-\frac{\lambda}{v}\int_{y}^{\infty} Q_B(x,y')dy'\right\} \tag{C.5}$$

where $T = 30/3600$h and $\tau_{AB} \geq 120/3600$h.

For the dataset after 1989/90,

$$p(x,y,A\to B) = \left(\frac{\lambda}{v}\right)^2 Q_B(x,y)\{Q_{A\cup B}(x,y+v\tau_{AB}) - Q_B(x,y+v\tau_{AB})\}$$
$$\times\exp\left[-\frac{\lambda}{v}\left\{\int_{y+v\tau_{AB}}^{\infty} Q_{A\cup B\setminus B}(x,y')dy' + \int_{y}^{\infty} Q_B(x,y')dy'\right\}\right] \tag{C.6}$$

where $\tau_{AB} > 60/3600$h.

6. $B \to A$

Same as $A \to B$ for exchanging the symbols A and B.

7. $C \to A$

For the dataset before 1988/89,

$$p(x,y,C\to A) = \frac{\lambda}{v}\left[\exp\left\{-\frac{\lambda}{v}\int_{y+v(\tau_{CA}+T)}^{\infty} Q_{A\cup B\cup C\setminus A\cup B}(x,y')dy'\right\} - \right.$$
$$\left.\exp\left\{-\frac{\lambda}{v}\int_{y+v(\tau_{CA}-T)}^{\infty} Q_{A\cup B\cup C\setminus A\cup B}(x,y')dy'\right\}\right]$$
$$\times\{Q_{A\cup B}(x,y) - Q_B(x,y)\}$$
$$\exp\left[-\frac{\lambda}{v}\left\{\int_{y}^{\infty} Q_{A\cup B}(x,y')dy' + \int_{0}^{y} Q_B(x,y')dy'\right\}\right] \tag{C.7}$$

where $T = 30/3600$h and $\tau_{CA} \geq 120/3600$h.

For the dataset after 1989/90,

$$p(x,y,C\to A) = \left(\frac{\lambda}{v}\right)^2\{Q_{A\cup B}(x,y) - Q_B(x,y)\}$$
$$\times\{Q_{A\cup B\cup C}(x,y+v\tau_{CA}) - Q_{A\cup B}(x,y+v\tau_{CA})\}$$
$$\times\exp\left\{-\frac{\lambda}{v}\int_{y+v\tau_{CA}}^{\infty} Q_{A\cup B\cup C\setminus A\cup B}(x,y')dy'\right\}$$
$$\times\exp\left[-\frac{\lambda}{v}\left\{\int_{y}^{\infty} Q_{A\cup B}(x,y')dy' + \int_{0}^{y} Q_B(x,y')dy'\right\}\right] \tag{C.8}$$

where $\tau_{CA} > 60/3600$h.

8. $C \to B$

Same as $C \to A$ for exchanging the symbols A and B.

C-2. CL mode

We have two platforms, top barrel and upper bridge, for CL mode. Once any observer on either platform detect the animal, the sighting is communicated to other observers by the researcher immediately. Hence, there are no duplicates in the CL mode. The detection function is given by

$$p(x, y, A \cup C) = \frac{\lambda}{v} Q_{A\cup C}(x, y) \exp\left\{-\frac{\lambda}{v}\int_y^{\infty} Q_{A\cup C}(x, y')dy'\right\}. \tag{C.9}$$

The contribution to the likelihood function of detection with each sighting pattern is calculated by the above probability density times the probability mass density of school size (Appendix C) divided by $\text{esw}_{A\cup C}$ when school sizes are confirmed. When school sizes are unconfirmed, the numerator is summed up over all school sizes. $\text{esw}_{A\cup C}$ is given by

$$\begin{aligned}\text{esw}_{A\cup C} = \sum_{s=1}^{\infty}\Bigg[\int_0^{x_{xmax}}\int_0^{\infty}&\frac{\lambda}{v} Q_{A\cup C}(x, y|s) \\ \times \exp\left\{-\frac{\lambda}{v}\int_y^{\infty} Q_{A\cup C}(x, y'|s)dy'\right\}& dxdy\Bigg]\pi(s),\end{aligned} \tag{C.10}$$

which is equal to Eq. (4) when k = CL mode.

Appendix D Scool size distribution

The probability mass function of true school size (s) is given by a negative binomial distribution for $s-1$,

$$\pi(s) = \frac{\Gamma(\phi_0 + s - 1)}{\Gamma(\phi_0)\Gamma(s)}\left(1 - \frac{\phi_0}{\phi_0 + \phi_1}\right)^{s-1}\left(\frac{\phi_0}{\phi_0 + \phi_1}\right)^{\phi_0}, \tag{D.1}$$

where $\phi_0 > 0$, $\phi_1 > 0$, and the parameter ϕ_1 is linked to the following covariates,

$\log(\phi_1) \sim$ Strata,

where Strata is categorical variables indicating the surveyed stratum. Note that $\mathrm{E}(s) = \phi_1 + 1$.

The probability mass function of unconfirmed school size given true school size is also given by a truncated negative binomial distribution,

$$\rho(z|s) = \frac{\Gamma(\psi_0 + s)}{\Gamma(\psi_0)\Gamma(s+1)[1 - \{\psi_0/(\psi_0 + \psi_1)\}^{\psi_0}]}\left(1 - \frac{\psi_0}{\psi_0 + \psi_1}\right)^{s}\left(\frac{\psi_0}{\psi_0 + \psi_1}\right)^{\psi_0}, \tag{D.2}$$

where $\psi_1 = \exp(\beta_0 + \beta_1 \times \log(s))s$, ψ_0 is a dispersion parameter, and $\mathrm{E}(z) = \psi_1$.

The probability of confirmation status c_k, which is given separately for each survey mode, is linked to the following covariates,

$\text{logit}(c_k) \sim \log(s) + \sqrt{x^2 + y^2}$ + Weather (for CL mode),

$\text{logit}(c_k) \sim \log(s) + x$ + Weather (for IO mode).

Appendix E The hazard probability model and the likelihood function

The detection probability density function of the animal positioned at the perpendicular distance x and the forward distance y assuming a Poisson surfacing pattern with the mean surfacing rate λ is

$$p(x,y) = \frac{\lambda}{v} Q(x,y) \exp\left\{ -\frac{\lambda}{v} \int_y^\infty Q(x,y')dy' \right\}, \tag{E.1}$$

where v is the vessel speed, and $Q(x,y)$ is a hazard probability function based on a logistic function (Appendix B). The parameters in Eq. (E1) are linked to various factors. Vessel, true school size, weather conditions (Beaufort or Sightability), and platforms were employed as the covariates for a detection process, while stratum and distance from the ice edge were linked with mean school size (Appendices B and C). We used Gaussian integration for all the integrals hereafter. In addition, we used Gaussian summation for $\sum_{s=1}^{\infty}$ (Monien 2006).

We construct a likelihood function conditioned on detection patterns and confirmation status of school size (Appendix C). For the confirmed school size, the likelihood function is

$$P_C(x_i, y_i, u_i, s_i) = \frac{c_k p_k(x_i, y_i, u_i|s_i)\pi(s_i)}{\text{esw}_k}, \tag{E.2}$$

and for the unconfirmed school size, the likelihood function is

$$P_U(x_i, y_i, u_i, z_i) = \frac{\sum_{s=1}^{\infty}(1-c_k)\rho(z_i|s)p_k(x_i, y_i, u_i|s)\pi(s)}{\text{esw}_k}, \tag{E.3}$$

where k is an index that denotes Passing/Closing mode, c_k is the probability of school size confirmation dependent on some covariates such as true school size, $\rho(z_i|s)$ is the probability that the school size is recorded as z_i given the true school size is s and the observed school size z_i is unconfirmed. u_i is a type of detection pattern, p_k is a detection probability density function given the mode k and the detection pattern u_i, and $\pi(s)$ is a probability mass function of true school size, and esw_k is

$$\text{esw}_k = \int_0^{x_{max}} \int_0^{\infty} \sum_{s=1}^{\infty} \sum_{u}^{\text{all patterns}} p_k(x, y, u|s)\pi(s)dxdy. \tag{E.4}$$

When the observed school size is eqaul to or greater than the certain threshold, the above probability for the confirmed and unconfirmed school size is modified to

$$P_C(x_i, y_i, u_i, s_i \geq s_{max}) = \frac{\sum_{s=s_{max}}^{\infty} c_k p_k(x_i, y_i, u_i|s)\pi(s)}{\text{esw}_k}, \tag{E.5}$$

$$P_U(x_i, y_i, u_i, z_i \geq z_{max}) = \frac{\sum_{s=1}^{\infty}(1-c_k)\{1 - \sum_{z=1}^{z_{max}-1} \rho(z|s)\}p_k(x_i, y_i, u_i|s)\pi(s)}{\text{esw}_k}. \tag{E.6}$$

where s_{max} and z_{max} is set to 9.

The mean value of true school size distribution, $\pi(s)$, is linked to the survey stratum (Appendix D). The confirmation probability is dependent on survey mode, weather condition, perpendicular distance (for IO mode) and radial distance (for CL mode) (Appendix D).

The total likelihood function is then given by

$$L = \prod_{i=1}^{n_C} P_C(x_i, y_i, u_i, s_i) \times \prod_{i=1}^{n_U} P_U(x_i, y_i, u_i, z_i), \tag{E.7}$$

where n_C and n_U are the numbers of the sightings with confirmed and unconfirmed school size, respectively. We estimate parameters by maximizing the logarithm of the total likelihood function. The maximization of the likelihood function is conducted separately for each circumpolar set.

Appendix F Abundance estimation

We use only the IO mode data in abundance estimation to circumvent possible biases that the CL mode data involve (Kishino and Kasamatsu, 1987; Branch and Butterworth, 2001), while we use both of the CL and IO mode data for parameter estimation as above mentioned. The population size is then estimated with a Horvitz-Thompson-like estimator,

$$\hat{P} = \frac{A}{2L} \sum_{i=1}^{n_P} \frac{\phi_1(\eta_i) + 1}{\sum_{s=1}^{\infty} \hat{\text{esw}}_{A \cup B \cup C}(s|\eta_i) \hat{\pi}(s|\eta_i)}, \tag{F.1}$$

where n_P is the number of the sightings in the IO mode, L is total survey distance, A is the size of survey area, η_i is a vector of covariates except for school sizes, and the numerator corresponds to the mean school size derived from a parametric distribution of school size (Appendix D).

An estimator for the unconditional asymptotic variance of $\hat{P}$ is then

$$\hat{\text{var}}(\hat{P}) = \left[\left\{ \frac{d\hat{P}(\theta)}{d\theta} \right\}^T I(\theta)^{-1} \frac{d\hat{P}(\theta)}{d\theta} \right]_{\theta=\hat{\theta}} + \frac{A^2}{J-1} \sum_{j=1}^{J} \frac{l_j}{L} (\hat{D}_j - D)^2, \tag{F.2}$$

where θ is a vector of estimated parameters, $I(\theta)$ is the Fisher information matrix obtained from the second derivative of the log-likelihood function that is often substituted by the Hessian matrix, and l_j ($j = 1, \ldots, J$; $\sum l_j = L$) is a replicated line. $\hat{D}_j$ is the density on replicate line j. If there is no sighting on replicate line j, $\hat{D}_j$ is defined as being equal to zero.

When the abundance estimates are obtained by strata, taking account of common estimated parameters across strata, the abundance estimate and its variance for the whole area are given by

$$\hat{P}_{\text{all strata}} = \sum_h A_h \hat{D}_h, \tag{F.3}$$

$$\hat{\text{var}}(\hat{P}_{\text{all strata}}) = \left[\left\{ \frac{d\hat{P}_{\text{all strata}}(\theta)}{d\theta} \right\}^T I(\theta)^{-1} \frac{d\hat{P}_{\text{all strata}}(\theta)}{d\theta} \right]_{\theta=\hat{\theta}} + \sum_h \frac{A_h^2}{J_h - 1} \sum_{j=1}^{J_h} \frac{l_{j,h}}{L_h} (\hat{D}_{j,h} - D_h)^2, \tag{F.4}$$

where the subscript h is the index of stratum.

The covariance between abundance estimates with different years taking account of common parameters is calculated by

$$\hat{\text{cov}}(\hat{P}_1, \hat{P}_2) = \left[\left\{ \frac{d\hat{P}_1(\theta)}{d\theta} \right\}^T I(\theta)^{-1} \frac{d\hat{P}_2(\theta)}{d\theta} \right]_{\theta=\hat{\theta}}, \tag{F.5}$$

where the subscripts denote different years and areas. The correlation matrix is obtained from the estimated variances and covariances. The additional variance is added to the estimated variances (Kitakado and Okamura, 2005, 2008, 2009). The final abundances for management areas were calculated using the estimates based on the additional variance blocks.

Table 1. Abundance estimates for each Management Area and the circumpolar estimates in CP II and CPIII.

OK preferred dataset

	Area	I	II	III	IV	V	VI	Total
CP II	Area size	433272	499193	484631	477030	963558	561763	3419447
	Abundance	**166,050**	**245,183**	**172,823**	**106,420**	**606,516**	**108,272**	**1,405,264**
	Density	0.383	0.491	0.357	0.223	0.629	0.193	0.411
	CV	0.199	0.157	0.205	0.177	0.173	0.267	0.118
CP III	Area size	677842	616407	752931	516276	834625	748269	4146350
	Abundance	**51,878**	**76,689**	**122,217**	**79,773**	**237,367**	**106,205**	**674,128**
	Density	0.077	0.124	0.162	0.155	0.284	0.142	0.163
	CV	0.120	0.185	0.152	0.322	0.115	0.147	0.088
Ratio of abund.		0.31	0.31	0.71	0.75	0.39	0.98	0.48
Ratio of dens.		0.20	0.25	0.46	0.69	0.45	0.74	0.40

Reference dataset

	Area	I	II	III	IV	V	VI	Total
CP II	Abundance	**160,205**	**242,002**	**162,637**	**101,310**	**585,216**	**98,839**	**1,350,210**
	Density	0.370	0.485	0.336	0.212	0.607	0.176	0.395
	CV	0.211	0.172	0.199	0.187	0.180	0.295	0.122
CP III	Abundance	**54,094**	**76,765**	**119,150**	**76,850**	**201,142**	**119,227**	**647,227**
	Density	0.080	0.125	0.158	0.149	0.241	0.159	0.156
	CV	0.154	0.199	0.170	0.326	0.131	0.168	0.100
Ratio of abund.		0.34	0.32	0.73	0.76	0.34	1.21	0.48
Ratio of dens.		0.22	0.26	0.47	0.70	0.40	0.91	0.40

Reference dataset (Confirmation treatment (2))

	Area	I	II	III	IV	V	VI	Total
CP II	Abundance	**137,005**	**212,225**	**140,551**	**87,372**	**507,703**	**85,021**	**1,169,877**
	Density	0.316	0.425	0.290	0.183	0.527	0.151	0.342
	CV	0.211	0.172	0.200	0.182	0.176	0.295	0.118
CP III	Abundance	**49,797**	**71,336**	**109,973**	**71,321**	**185,171**	**110,358**	**597,957**
	Density	0.073	0.116	0.146	0.138	0.222	0.147	0.144
	CV	0.154	0.198	0.169	0.329	0.130	0.168	0.100
Ratio of abund.		0.36	0.34	0.78	0.82	0.36	1.30	0.51
Ratio of dens.		0.23	0.27	0.50	0.75	0.42	0.97	0.42

Reference dataset (Confirmation treatment (1))

	Area	I	II	III	IV	V	VI	Total
CP II	Abundance	**142,722**	**219,569**	**146,021**	**90,707**	**526,852**	**88,368**	**1,214,240**
	Density	0.329	0.440	0.301	0.190	0.547	0.157	0.355
	CV	0.210	0.171	0.199	0.182	0.176	0.295	0.117
CP III	Abundance	**52,869**	**75,063**	**116,137**	**75,271**	**196,136**	**116,439**	**631,915**
	Density	0.078	0.122	0.154	0.146	0.235	0.156	0.152
	CV	0.153	0.198	0.169	0.325	0.130	0.167	0.098
Ratio of abund.		0.37	0.34	0.80	0.83	0.37	1.32	0.52
Ratio of dens.		0.24	0.28	0.51	0.77	0.43	0.99	0.43

Reference dataset (SSX data)

	Area	I	II	III	IV	V	VI	Total
CP II	Abundance	**138,696**	**213,647**	**142,275**	**88,023**	**511,654**	**85,286**	**1,179,581**
	Density	0.320	0.428	0.294	0.185	0.531	0.152	0.345
	CV	0.209	0.170	0.198	0.181	0.175	0.293	0.116
CP III	Abundance	**52,313**	**74,130**	**115,439**	**74,389**	**195,029**	**115,368**	**626,666**
	Density	0.077	0.120	0.153	0.144	0.234	0.154	0.151
	CV	0.148	0.192	0.165	0.324	0.127	0.174	0.092
Ratio of abund.		0.38	0.35	0.81	0.85	0.38	1.35	0.53
Ratio of dens.		0.24	0.28	0.52	0.78	0.44	1.02	0.44

Table 2. Abundance estimates for each Management Area and the circumpolar estimates in CP II and CPIII. Additional sensitivity analyses.

Reference dataset (ss error = 0)

	Area	I	II	III	IV	V	VI	Total
CP II	Abundance	**89,275**	**145,650**	**93,100**	**59,579**	**331,066**	**53,766**	**772,437**
	Density	0.206	0.292	0.192	0.125	0.344	0.096	0.226
	CV	0.213	0.175	0.199	0.170	0.162	0.287	0.102
CP III	Abundance	**39,870**	**58,039**	**88,945**	**57,221**	**145,452**	**89,927**	**479,455**
	Density	0.059	0.094	0.118	0.111	0.174	0.120	0.116
	CV	0.153	0.198	0.166	0.338	0.128	0.167	0.097
Ratio of abund.		0.45	0.40	0.96	0.96	0.44	1.67	0.62
Ratio of dens.		0.29	0.32	0.61	0.89	0.51	1.26	0.51

Preferred dataset (including "BH")

	Area	I	II	III	IV	V	VI	Total
CP II	Abundance	**166,050**	**245,183**	**172,823**	**106,420**	**606,516**	**108,272**	**1,405,264**
	Density	0.383	0.491	0.357	0.223	0.629	0.193	0.411
	CV	0.199	0.157	0.205	0.177	0.173	0.267	0.118
CP III	Abundance	**51,973**	**76,831**	**122,437**	**75,271**	**237,744**	**106,397**	**670,652**
	Density	0.077	0.125	0.163	0.146	0.285	0.142	0.162
	CV	0.120	0.185	0.152	0.325	0.115	0.147	0.088
Ratio of abund.		0.31	0.31	0.71	0.71	0.39	0.98	0.48
Ratio of dens.		0.20	0.25	0.46	0.65	0.45	0.74	0.39

Reference dataset (special handling for ss = 1)

	Area	I	II	III	IV	V	VI	Total
CP II	Abundance	**187,780**	**267,645**	**187,058**	**119,754**	**669,673**	**119,404**	**1,551,314**
	Density	0.433	0.536	0.386	0.251	0.695	0.213	0.454
	CV	0.232	0.185	0.217	0.215	0.204	0.322	0.116
CP III	Abundance	**55,283**	**78,423**	**121,271**	**78,546**	**203,724**	**121,313**	**658,560**
	Density	0.082	0.127	0.161	0.152	0.244	0.162	0.159
	CV	0.156	0.200	0.170	0.325	0.133	0.168	0.097
Ratio of abund.		0.29	0.29	0.65	0.66	0.30	1.02	0.42
Ratio of dens.		0.19	0.24	0.42	0.61	0.35	0.76	0.35

Preferred dataset (including possible duplicates)

	Area	I	II	III	IV	V	VI	Total
CP II	Abundance	**151,529**	**228,105**	**158,752**	**96,052**	**550,794**	**99,218**	**1,284,450**
	Density	0.350	0.457	0.328	0.201	0.572	0.177	0.376
	CV	0.194	0.154	0.204	0.172	0.166	0.262	0.112
CP III	Abundance	**48,712**	**71,363**	**114,300**	**73,670**	**225,557**	**98,933**	**632,536**
	Density	0.072	0.116	0.152	0.143	0.270	0.132	0.153
	CV	0.117	0.184	0.151	0.313	0.114	0.142	0.085
Ratio of abund.		0.32	0.31	0.72	0.77	0.41	1.00	0.49
Ratio of dens.		0.21	0.25	0.46	0.71	0.47	0.75	0.41